

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-259316
 (43)Date of publication of application : 24.09.1999

(51)Int.CI. G06F 9/46
 G06F 9/46

(21)Application number : 11-002585 (71)Applicant : HITACHI LTD
 (22)Date of filing : 08.01.1999 (72)Inventor : IMADA TOYOHISA
 HOSHINA TAKUICHI

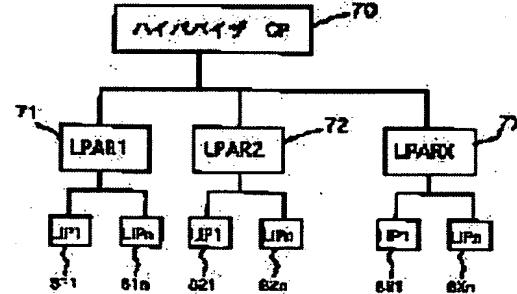
(30)Priority
 Priority number : 10 2795 Priority date : 09.01.1998 Priority country : JP

(54) CPU CAPABILITY ADJUSTMENT METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To effectively utilize for operation management, etc., by allocating a differential capability between the CPU capability of a physical computer and total CPU capabilities to a virtual computer for maintenance.

SOLUTION: A hyper visor 70 prepares an LPARX (7X) for maintenance apart from LPAR groups (71 and 72) used by a user in initial processing and automatically sets main storage for the LPARX and logic IP. When a CPU capability coefficient is designated, a capability other than the designated CPU capability is allocated to the LPARX and activated. After that, LPARs 1 and 2 are activated and three LPARs are activated as the entire system. Main storage, CPU capability, etc., which are designated by the user are allocated to the LPARs 1 and 2 respectively. CPU capabilities of the LPARs 1 and 2 are designated with a ratio value (real service rate) of the entire system in accordance with needed capability value. The user designates only ratio value (user definition service rate) that corresponds to value designated as a needed CPU capability.



LEGAL STATUS

[Date of request for examination] 19.03.2003

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平11-259316

(43)公開日 平成11年(1999)9月24日

(51)Int.Cl.⁶
G 0 6 F 9/46

識別記号
3 5 0
3 4 0

F I
G 0 6 F 9/46

3 5 0
3 4 0 D

審査請求 未請求 請求項の数 7 O.L. (全 6 頁)

(21)出願番号 特願平11-2585
(22)出願日 平成11年(1999)1月8日
(31)優先権主張番号 特願平10-2795
(32)優先日 平10(1998)1月9日
(33)優先権主張国 日本 (JP)

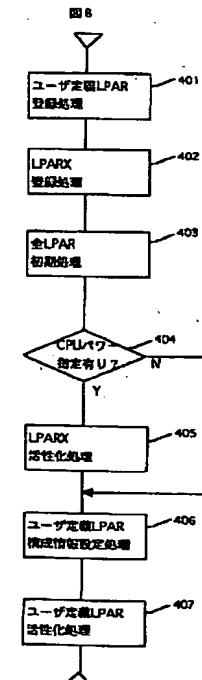
(71)出願人 000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目6番地
(72)発明者 今田 豊寿
神奈川県秦野市堀山下1番地 株式会社日立製作所汎用コンピュータ事業部内
(72)発明者 星名 卓一
神奈川県秦野市堀山下1番地 株式会社日立製作所汎用コンピュータ事業部内
(74)代理人 弁理士 小川 勝男

(54)【発明の名称】 C P U能力調整方法

(57)【要約】

【課題】 ユーザが必要としているトータルC P U能力は導入された計算機のC P U能力よりは小さく、その差分の無駄になっているC P U能力を有効に利用するシステムを提供する。

【解決手段】 物理計算機上で複数の仮想計算機が動作する仮想計算機システムにおいて、各仮想計算機が必要とするC P U能力を合計したトータルC P U能力を前記物理計算機のC P U能力内で任意に設定し、各仮想計算機を前記設定されたトータルC P U能力の範囲内で動作させるようにC P U使用率を決定し、前記物理計算機のC P U能力と前記トータルC P U能力の差分のC P U能力を保守用仮想計算機に割り当てる。



【特許請求の範囲】

【請求項1】 物理計算機上で複数の仮想計算機が動作する仮想計算機システムにおいて、各仮想計算機が必要とするCPU能力を合計したトータルCPU能力を前記物理計算機のCPU能力内で任意に設定するステップと、各仮想計算機を前記設定されたトータルCPU能力の範囲内で動作させるようにCPU使用率を決定するステップと、前記物理計算機のCPU能力と前記トータルCPU能力の差分のCPU能力を保守用仮想計算機に割り当てるステップを有することを特徴とするCPU能力調整方法。

【請求項2】 物理計算機上で複数の仮想計算機が動作する仮想計算機システムにおいて、少なくとも1つの第1の仮想計算機を登録するステップと、第2の仮想計算機を登録するステップと、前記第1の仮想計算機及び前記第2の仮想計算機を初期化するステップと、第1のCPU使用率に基づいて前記第2の仮想計算機を活性化するステップと、前記第1のCPU使用率に関連する第2のCPU使用率で前記第1の仮想計算機を活性化するステップとを有することを特徴とするCPU能力調整方法。

【請求項3】 前記物理計算機はサービスプロセッサを有し、該サービスプロセッサから前記第1のCPU使用率を設定するステップとを有することを特徴とする請求項2記載のCPU能力調整方法。

【請求項4】 前記第1の仮想計算機はユーザが定義する仮想計算機であることを特徴とする請求項2記載のCPU能力調整方法。

【請求項5】 前記第2の仮想計算機はシステムの保守用の仮想計算機であることを特徴とする請求項2記載のCPU能力調整方法。

【請求項6】 物理計算機上で複数の仮想計算機が動作する仮想計算機システムにおいて、各仮想計算機が必要とするCPU能力を合計するステップと、前記物理計算機のCPU能力と前記トータルCPU能力の差分を計算するステップと、前記差分のCPU能力を持つ仮想計算機を生成するステップとを有することを特徴とするCPU能力調整方法。

【請求項7】 物理計算機上で複数の仮想計算機が動作する仮想計算機システムであって、第1のCPU使用率を割り当てられるシステム保守用仮想計算機と、前記第1のCPU使用率に応じて決定される第2のCPU使用率を割り当てられる少なくとも1つのユーザ定義仮想計算機とを有することを特徴とする仮想計算機システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、計算機システムのCPU資源を効率的に使用する方法に関し、特に仮想計算機システムにおいて無駄になっているCPU資源を効率的に利用する方法に関する。

【0002】

【従来の技術】 従来、計算機ユーザは、導入した計算機(物理計算機)上に複数の仮想計算機から構成される仮想計算機システムを構築する際、各仮想計算機が必要とするCPU能力を合計してトータルCPU能力を求め、CPU能力の大きさにより段階的に設定されている製品モデルの中から、そのトータルCPU能力を満足する製品モデルを決めている。そして、その決定された製品モデルのCPU能力に応じて各仮想計算機にサービスする割合(サービス率)を指定していた。各仮想計算機にサービス率を指定する機能に関しては特開平9-81401に開示されている。

【0003】

【発明が解決しようとする課題】 即ち、ユーザが必要としているトータルCPU能力は導入された計算機のCPU能力よりも小さく、その差分のCPU能力が無駄になっていた。

【0004】 本願発明の目的は、この無駄になっているCPU能力を、システム全体のモニタリング又は保守用として動作する仮想計算機に割り当て、稼働管理などに効率的に利用することにある。

【0005】

【課題を解決するための手段】 本発明は、物理計算機上で複数の仮想計算機が動作する仮想計算機システムにおいて、各仮想計算機が必要とするCPU能力を合計したトータルCPU能力を前記物理計算機のCPU能力内で任意に設定し、各仮想計算機を前記設定されたトータルCPU能力の範囲内で動作させるようにCPU使用率を決定し、前記物理計算機のCPU能力と前記トータルCPU能力の差分のCPU能力を保守用仮想計算機に割り当てる。

【0006】 又、本発明は、物理計算機上で複数の仮想計算機が動作する仮想計算機システムにおいて、各仮想計算機が必要とするCPU能力を合計し、前記物理計算機のCPU能力と前記トータルCPU能力の差分を計算し、前記差分のCPU能力を持つ仮想計算機を生成する。

【0007】

【発明の実施の形態】 以下、本発明の実施例を図面により詳細に説明する。

【0008】 図1は、本発明が適用される物理計算機システムの構成を示す。10、11、…、1nはそれぞれ物理プロセッサIPO、物理プロセッサIPI、…、物理プロセッサIPnを示している。20は主記憶装置を示している。30、31、…、3nはそれぞれI/OプロセッサIOP0、I/OプロセッサIOP1、…、I/OプロセッサIOPnを示している。40はサービスプロセッサ(SVP)、50はコンソールディスプレイ装置(CD)、60はハードディスク装置(HD)を示している。70は仮想計算機システ

ム全体の制御プログラム(ハイパバイザ)を示している。

【0009】図2は本発明に係る仮想計算機システムの構成を示す。70はハイパバイザ、71、72、7Xはそれぞれ仮想計算機(以後、LPARと呼ぶ)LPAR1、LPAR2、LPARXを表す。LPAR1、LPAR2はユーザが定義した仮想計算機であり、LPARXは本発明で新たに設けた保守用の仮想計算機である。811、…、81nはLPAR1制御下のn台の論理プロセッサ(以後、LIPと呼ぶ)LIP1、…、LIPnを表す。以下同様にして、821、…、82nはLPAR2制御下のn台の論理プロセッサLIP1、…、LIPnを表す。8X1、…、8XnはLPARX制御下のn台の論理プロセッサLIP1、…、LIPnを表す。ハイパバイザ70は主記憶装置20内に存在する。尚、本実施例ではユーザが定義した仮想計算機を2台としているが、台数はこれに限定されるものではない。

【0010】図3は、ハイパバイザ70の構成を示す。ハイパバイザ70は各LPARのスケジュールを行うスケジューラ90、各LPARの初期処理及びLPARXの活性化を行う初期処理部91、SVPから供給される物理資源及び各LPARにおける論理資源を管理する資源管理部92、各LPARの論理プロセッサの実行を制御する論理プロセッサ制御部93、各LPARの資源割り当て、各LPARの稼働状態等の画面表示を制御するフレーム制御部94等から構成される。

【0011】図4は主記憶装置20とハードディスクHD60のデータの格納状態を示す。ハードディスクHD60にはハイパバイザ格納部61、I/O構成情報部62、モニタプログラム格納部63が設けられている。ハイパバイザの起動時に、ハイパバイザ格納部61に格納されているハイパバイザが主記憶装置20の拡張ハードウェアシステムエリア(EHSA)22の該当部(23)にローディングされる。モニタプログラム格納部63に格納されたLPARX用のモニタプログラムも、ハイパバイザの起動時にEHSA22の該当部(24)にローディングされる。21はハードウェアシステムエリア(HSA)であり、計算機の立ち上げ時にハードウェアの制御情報が格納される。24はLPAR情報テーブルであり、各LPARの構成情報が格納される。

【0012】図5はLPAR情報テーブル100の詳細を示す。101はLPAR名称欄である。102は主記憶容量設定欄であり、各LPARに与えられる主記憶容量を示す。103は論理IP定義欄であり、各LPARで生成される論理IPを示す。104はユーザ定義のCPUサービス率設定欄である。105は実際に各LPARに割り当たられるCPUサービス率(実サービス率)を示す。

【0013】図6は本発明で新たに設けたSVPのCP

U能力制御フレームを示す。CPU能力制御フレームは、ユーザの必要CPU能力を、ユーザが導入した計算機システムのCPU能力を100とした場合の係数で指定する。例えば、当計算機システムのCPU能力を100MIPSとし、ユーザが使用する仮想計算機のLPAR1及びLPAR2のCPU能力がそれぞれA1MIPS、A2MIPSでA1とA2の合計を80MIPSとする。この場合、80という値をCPU能力係数として指定する。前記CPU能力制御フレームで指定されたCPU能力係数はCPU能力指定テーブル(図7)に格納され、ハイパバイザ70の資源管理部92により管理される。

【0014】次に、本実施例におけるCPU能力の制御方法を説明する。

【0015】まず、物理計算機を仮想計算機モードに指定してシステムを立ち上げる。システムを立ち上げると、ハイパバイザがハードディスク装置(HD)から拡張ハードウェアシステムエリア(EHSA)にローディングされ、起動される。

【0016】ハイパバイザは初期処理にて、ユーザが使用するLPAR群とは別に保守用のLPAR(LPARX)を準備し、LPARX用の主記憶及び論理IPを自動設定する。CPU能力係数が指定されている時は、その指定されたCPU能力以外の能力(上記の例の場合は20)をLPARXに割り当てる。その後、LPAR1及びLPAR2を活性化させ、システム全体として、3LPARを活性化させる。LPAR1及びLPAR2にはそれぞれユーザ指定された主記憶及びCPU能力等を割り当てる。LPAR1及びLPAR2のCPU能力は上記必要能力値(A1、A2)に応じて、システム全体の割合値(実サービス率)で指定する。尚、ユーザは必要CPU能力として指定した値(A1、A2)に応じた割合値(ユーザ定義サービス率)だけを指定する。

【0017】次に、システム全体のCPU能力の各LPARへの割り当てる(サービス率)について、具体的に説明する。従来(CPU能力指定機能がない場合)は、各LPARの必要CPU能力A1、A2に対して、例えばA1:A2=60:40の場合、LPARフレームでの設定に従い、ハイパバイザはLPAR1、LPAR2へはシステム全体のCPU能力の60%、40%を割り当てる。これに対して、本発明では、システム全体のCPU能力の80%をLPAR1、LPAR2に割り当てる。具体的には、LPAR1に対してはシステム全体のCPU能力の48%(60%×0.08)、LPAR2に対してはシステム全体のCPU能力の32%(40%×0.08)を割り当てる。

【0018】尚、ユーザ定義LPARのCPUサービス率に変動が生じないようにするため、LPARXはその

CPU能力を一定(本実施例では20%)に保つ必要がある。このために、LPARXに対して、CPUサービス率の上限を保証するリソースキャップ機能(RC機能)を設定する。RC機能とは、あるインターバル中ににおいてCPUサービス率の指定上限値以上のCPUサービスを抑止する機能である。さらに、CPUサービス率の下限を保証するウェイトコンプリഷョン機能(WC機能)を設定する。WC機能はプログラムがWAITしても制御をH/Wが保留したまま保持する機能であり、WAITする頻度が高い場合でも、CPUをCPUサービス率の指定値まで使用できる機能である。RC機能及びWC機能は公知の機能であり、特開平9-81401等に開示されている。以上のように、RC機能及びWC機能を設定することにより、LPARXは常に一定のCPUサービス率を保持できる。

【0019】最後に、ハイパバイザの概略処理フローを図8により説明する。図8はハイパバイザ初期処理の内、ユーザ定義LPARの登録処理からユーザ定義LPARの活性化処理までを示している。ユーザ定義LPAR登録処理ステップ401では、ハードディスク装置HD60のI/O構成情報62から指定したLPAR名称を取り出し、ハイパバイザの資源管理部が管理するLPAR情報テーブル100に登録する。LPARX登録処理ステップ402では、本発明で新たに設けるLPARをLPARXという名称で登録する。又、LPARXに割り当てる主記憶エリア及び論理IPの定義情報をLPAR情報テーブル100に登録する。この時、論理IPは物理IPの数と同じ数だけ登録する。この論理IPは共有モード(他LPARと物理IPを共有して使用するモード)とする。全LPAR初期処理ステップ403では、LPAR情報テーブル100に登録されたLPARの初期処理を行う。具体的には各LPARに対応する論理IPの生成及び論理IPの内部テーブルの初期化を行う。CPU能力指定判定ステップ404では、資源管理部92が管理している能力指定値テーブル300を読み出し、CPU能力指定がある場合(CPU能力係数が100でない場合)、LPARXの活性化処理ステップ405を実行する。指定がない場合(CPU能力係数が100の場合)はLPARXの活性化処理は行わない。LPARX活性化処理ステップ405では、LPARXのアクティベイト処理を行う。具体的には、LPARXに割り当てられた主記憶、論理IP等をオンラインにする処理であり、LPAR情報テーブル100に設定された構成情報を用いてオンライン処理を行う。その後、モニタプログラムを自動IPLする。以上でハイパバイザ初期処理が完了する。これ以降、LPARフレームにて、ユーザ定義LPARの構成情報設定処理406が可能となる。構成設定後、ユーザ定義LPARの活性化処理407を行い、前述したCPU能力係数をふまえたCPUサービス率を割り当てる。その後、ユーザ定義LPAR

毎にOSをIPLした後、実稼働にはいる。尚、LPARXにて動作しているモニタプログラムはLPARフレームにて、ユーザ定義LPARの稼働状態を見る事もできる。

【0020】動的にCPU能力係数を変更する場合、ハイパバイザはSVPから状態変更通知を受け取るとその情報を読み出し、資源管理部92の能力指定値テーブル300を変更する。その後、ハイパバイザは全LPARに対して、前述したCPUの割り当て方法と同様にシステム全体における割り当て値を決め、各LPARのサービス率を決定する。動的にCPU能力係数を変更する処理を行うために、SVPからハイパバイザへの連絡通知を行うインターフェース(外部割込)を使用する。即ち、SVPとハイパバイザ間のやりとりを司るHVA命令にCPU能力係数データを読み込む機能を追加することにより、割込時にCPU能力係数をハイパバイザへ連絡する。動的に各LPARのCPU割り当て値を変更する処理は、LPARフレームを使用することにより実現できる。

【0021】図9は、従来技術のCPU使用率(サービス率)及び本発明によるCPU使用率(サービス率)を比較したものである。LPAR1、LPAR2については、必要CPU能力A1、A2に対応するCPU使用率 α_1 、 α_2 が、本発明により変化しないことを示している。即ち、従来は、必要CPU能力A1、A2以外のCPU能力(100 - α_1 - α_2)をイベントサーチが使用していたが、本発明によりこのCPU能力(100 - α_1 - α_2)がLPARXに割り当られる。

【0022】

【発明の効果】本発明によれば、仮想計算機システムにおいて、ユーザは必要なCPU能力のみを使用することが可能であり、ユーザに不要なCPU能力を、保守用仮想計算機に割り当てることで、システム稼働管理等ができるようになり、CPU資源の有効活用ができる。

【図面の簡単な説明】

【図1】本発明が適用される物理計算機システムの構成図である。

【図2】本発明に係る仮想計算機システムの構成図である。

【図3】図2におけるハイパバイザの構成を示す図である。

【図4】図1における主記憶装置とハードディスクHDのデータの格納状態を示す図である。

【図5】図4におけるLPAR情報テーブルの詳細を示す図である。

【図6】本発明に係るCPU能力制御フレームを示す図である。

【図7】本発明に係るCPU能力指定テーブルを示す図である。

【図8】図2におけるハイパバイザの概略処理フローを

示す図である。

【図9】従来技術及び本発明におけるCPU使用率(サービス率)を示す表である。

【符号の説明】

10、11、…、1n 物理プロセッサ(I P)

20 主記憶装置

30、31、…、3n I /Oプロセッサ(I O P)

40 サービスプロセッサ(S VP)

50 コンソールディスプレイ装置(C D)

60 ハードディスク装置(H D)

70 ハイパバイザ(CP)

71、72、7n 仮想計算機(L P A R)

811、…、81n 論理プロセッサ(L I P)

90 スケジューラ

91 初期処理部

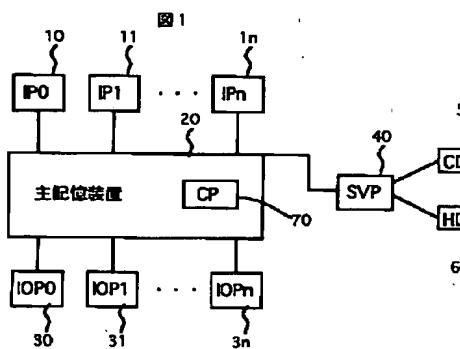
92 資源管理部

93 論理プロセッサ制御部

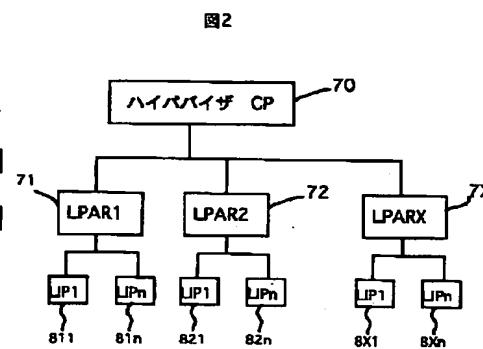
94 フレーム制御部

32 200 CPU能力制御フレーム

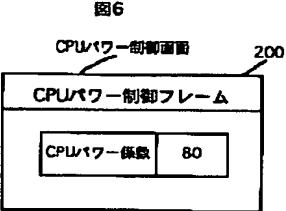
【図1】



【図2】

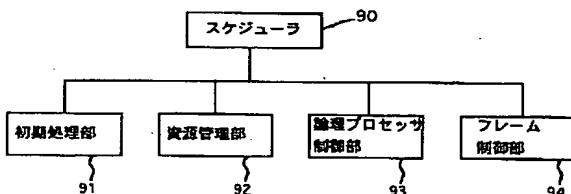


【図6】



【図3】

図3



【図5】

図4

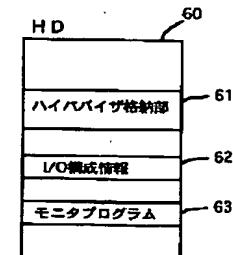
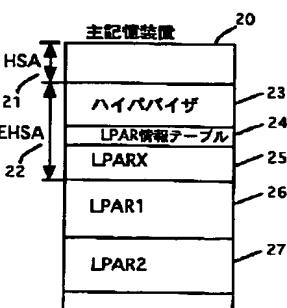


図5

LPAR名称	主記憶容量	論理IP定義	ユーザ定義サービス率	実サービス率
LPAR1	m1	IP1,IP2	60	48
LPAR2	m2	IP1,IP2	40	32
LPARX	mX	IP1,IP2	—	20

【図7】

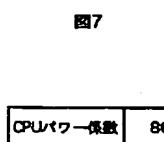


図7

【図9】

CPU処理項目	CPU使用率(%)	
	従来技術	本発明
LPAR1	α_1	α_1
LPAR2	α_2	α_2
イベントサーチ	$100 - \alpha_1 - \alpha_2$	ほとんど0
LPARX	—	$100 - \alpha_1 - \alpha_2$
システム全体	100	100

【 図8 】

図8

